DOCUMENT RESUME

ED 325 512 TM 015 750

AUTHOR Moore, Alan D.; Cross, Tracy L.
TITLE The Spirit Is Willing, but the Flesh Is Weak: Criterion-Referenced Testing in Wyoming.
PUB DATE Oct 90
NOTE 14p.; Paper presented at the Annual Meeting of the Northern Rocky Mountain Educational Research Association (8th, Greeley, CO, October 4-6, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Administrator Attitudes; *Criterion Referenced Tests; *Educational Needs; Elementary Secondary Education; Norm Referenced Tests; Public Schools; Questionnaires; School Districts; *School Surveys; State Programs; State Surveys; *Superintendents; *Testing Problems; Testing Programs; Test Use
IDENTIFIERS *Wyoming

ABSTRACT

The perceived needs of public school personnel in Wyoming with respect to the development and use of criterion-referenced tests (CRTs) as part of a district-wide testing program were assessed using a survey designed for the study. Questionnaires were sent to all superintendents, assistant superintendents, and curriculum directors in the state. Of the original 67 questionnaires mailed, 51 (76%) were returned after a follow-up mailing to non-respondents from the first mailing. Data were coded, and frequencies were tabulated for each item on the questionnaire. In all, 42% of the respondents reported that their testing program was primarily norm-referenced. Another 42% reported that a combination of norm-referenced tests and CRTs was in use. Over half (53%) of the respondents reported great or extreme need, 27% reported moderate need, and 16% reported slight need or no need for professional development in constructing, validating, and using CRTs. On-site consulting, professional development seminars, extension course work through the University of Wyoming (Laramie), and a summer institute were seen as most desirable or feasible. Results indicate that Wyoming schools are in a state of transition in terms of criterion-referenced testing; these schools are willing to adopt CRTs, but are too weak in skills to use them effectively. A list of 16 skills that educators need in order to use CRTs appropriately is included. Five tables contain information about survey responses. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED325512

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
ALAN D. MOORE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

THE SPIRIT IS WILLING, BUT THE FLESH IS WEAK: CRITERION-REFERENCED TESTING IN WYOMING

Alan D. Moore

and

Tracy L. Cross

University of Wyoming

ABSTRACT

As interest in criterion-referenced testing in the public schools increases, many districts lack the resources and expertise to develop these tests. This study assessed the perceived needs of public school personnel in the state of Wyoming with respect to the development and use of criterion-referenced tests as part of district-wide testing programs. Although districts were in various stages in the development and use of these tests, a sizeable need for technical support and training of district personnel was found to exist. The most desirable way of delivering this support was found to be through on-site consulting and professional development seminars.

Paper presented at the 8th Annual Meeting of the Northern Rocky Mountain Educational Research Association, October 4-6, 1990, Greeley, Colorado

TM015750

How can we improve schools? How can we insure equality of educational opportunity for all students? How can our students be prepared for an ever-changing Information Age society? How can we prepare and support effective teachers? These are core questions which occupy the thoughts of professional educators. Answers to these questions provide educators with goals. When programs are implemented to meet these goals, a link between goals and reality is provided. However, the other critical link between goals and reality is provided by meaningful educational measurement. It is the systematic measurement process which informs us of the successes and failures in meeting our educational goals. Unfortunately, knowing <u>that</u> we should provide good educational measurement programs is not the same as <u>being able to</u> do so. In this sense, "the spirit is willing, but the flesh is weak."

Several recent trends related to educational measurement and testing currently converge on school districts nation-wide including those in Wyoming. Popham and Hambleton (1990) warn that with advent of minimum competency testing, use of tests for evaluation of teachers, schools, and districts, National Teachers Examinations, outcome-based educational measurement has increased in importance in the last decade. As the demand for accountability in public education has increased, many school districts in Wyoming are being asked to reexamine their district-wide testing programs. Also, schools which decide to adopt goals of Mastery Learning, the Effective Schools Movement, School Incentive Awards Program, and Outcome-based Education may discover quickly that their testing and

assessment programs are inadequate for the information needs of the district. At the same time, a better understanding of the role of norm-referenced tests, what they can and cannot be expected to do, has created the demand and expectation that more criterion-referenced testing should be done in schools. In his 1987 Presidential Address to the National Council of Measurement in Education, Richard Jaeger reviewed the history of the past 20 years in educational measurement. In the article following that address he writes that "since 1967, criterion-referenced measurement has spawned a minor industry" and that "practitioners have become criterion-referenced zealots." (Jaeger, 1987, p. 7).

This trend toward more widespread use of criterion-referenced tests as part of testing programs has made it important for districts to either purchase these tests or develop them using district personnel. And since these tests are targeted toward local and district instructional objectives, few commercially published criterion-referenced tests will be found adequate for the measurement needs of individual districts. Who is to guide the development of these tests? Without expertise in developing these tests, we are in danger of losing the strongest feature of commercially available standardized tests - the high quality of their construction and validation.

Since much of the discussion of criterion-referenced tests has occurred only recently, many school administrators lack the appropriate training in the development and use of these tests. Locating administrators with this training in educational

measurement is difficult. This is particularly a problem in Wyoming since school districts are generally too small to support specialists in educational testing and measurement. And teachers rarely have the training to undertake this leadership. During pre-service training for teachers many states do not require even a single course in educational testing. Although the College of Education at the University of Wyoming currently requires such a course, this is likely to be the only training new teachers and new administrators receive in completing baccalaureate or advanced degree work. Even those with training in norm-referenced testing and educational measurement may not have studied methods of constructing criterion-referenced tests.

There are examples in Wyoming of districts which have begun to implement successful testing programs which include criterion-referenced testing. Campbell County School District #1 in Gillette has implemented a district-wide testing program using both norm- and criterion-referenced tests. (Weber, 1989). A norm-referenced test is administered periodically to all students in grades 1-12 to provide extra-district comparison information. A criterion-referenced testing program is maintained to assess student mastery of district objectives. A district item bank is maintained which enables teachers to construct their own criterion-referenced tests and have them scored and recorded using an integrated computerized system.

Sweetwater County District #1, Rock Springs, has implemented the Comprehensive Assessment Program. (Sheinker, 1990). A norm-

referenced test, the Survey of Basic Skills is administered annually and used for screening individual students for further assessment and for external comparison. Locally-developed Mastery Tests are administered as pretests and post-tests to students in the Fall and Spring. These are criterion-referenced tests with immediate results provided to teachers for their use in making instructional decisions.

But how widespread are these efforts to initiate and support systematic educational measurement in the schools of Wyoming? In this study, we sought to determine the degree of need in the state among public school personnel for training and technical support.

Methodology

A survey was designed to use in assessing the needs of public school districts in Wyoming with respect to the development and use of criterion-referenced tests in district-wide testing programs. Questionnaires were sent to all superintendents, assistant superintendents, and curriculum directors in the state. The cover letter asked that the person most responsible for testing in the district complete and return the questionnaire. Of the original 67 questionnaires mailed, 51 (76%) were returned after a follow-up mailing to nonrespondents from the first mailing. Data were coded and frequencies tabulated for each item on the questionnaire.

Results

Those responsible for test coordination in most districts include superintendents, assistant superintendents, guidance counselors, and curriculum directors. (See Table 1.) These

personnel represent 63% of the sample. Over half (54%) of the respondents reported five percent or less of the person's position is directly related to test coordination, almost three fourths (74%) reported 10 percent or less, and apparently only two people in the state are responsible for more than half-time for testing. (See Table 2.)

About two fifths (42%) reported that their testing program was primarily norm-referenced, and another two fifths (42%) reported a combination of norm-referenced and criterion-referenced testing. Only four respondents reported exclusive use of criterion-referenced tests district-wide.

Nearly two-thirds (64%) of the tests used are commercially available tests, but a sizable percentage (38%) reported the use of tests developed locally by school personnel.

In Table 3 are displayed the results concerning stages of development and use of criterion-referenced tests. The maximum use of criterion-referenced tests was reported to be in math (26%), reading (22%), and the other language arts (20%). They are used less in science (18%) and social studies (14%). Between one sixth and one fifth of the respondents reported they were in the stages of planning and/or development of objectives for criterion-referenced test in all these areas. However, at most 5 respondents (10%) reported that they were in the stage of writing items for tests areas other than math, where 8 respondents (16%) reported being in the item-writing stage.

Over half (53%) of the respondents reported great or extreme need of professional development in constructing, validating and using criterion-referenced tests, 27% reported moderate need and only 16% reported slight need or no need. (See Table 4). The kinds of professional development seen as most desirable and feasible were individual on-site consulting, professional development seminars, extension course work through UW, and a summer institute. Computer support was listed by a sizeable minority (29%) while very few expressed a need for individual telephone consulting and course work at UW. (See Table 5).

Discussion

The results of this study indicate that school districts in Wyoming are in a transitional period concerning criterion-referenced testing. Most districts have not allocated personnel resources for standardized testing. Norm-referenced tests or norm-referenced tests with one or two criterion-referenced tests are predominant. A minority of districts are in some stage of development and use of these tests, primarily in mathematics, and to a lesser extent in reading and language arts. Respondents to this survey expressed a critical need for district support in the developing, using, and revising criterion-referenced tests. The most desirable and feasible delivery system for this support is seen to be individual consultation with districts, professional development seminars, and extension course work in the district.

A beginning set of objectives for helping school districts to make better use of criterion-referenced testing would be the set of

"minimum competencies" for school administrators in the area of testing and measurement recently identified by Popham and Hambleton (1990).

Those responsible for district testing programs should be able to:

1) know basic measurement terminology
2) know similarities and differences between norm-referenced and criterion-referenced test development and test-score interpretation
3) describe steps in constructing criterion-referenced tests
4) design and carry out appropriate reliability and validity studies for locally constructed criterion-referenced tests
5) use basic statistical information in interpreting group and individual test scores
6) identify standard-setting issues and methods to support assessment-based decisions
7) identify potentially useful commercially available norm- and criterion-referenced tests to meet specific information needs
8) use appropriate evaluation criteria to choose among potential norm- and criterion-referenced tests
9) write, review, and use test-item specifications
10) identify strengths and weaknesses of various item formats
11) use multiple types of formats in preparing test items
12) use standard judgmental and empirical techniques for

evaluating performance test items

13) assemble tests including final item selection

14) identify the merits and demerits of major test-score reporting scales

15) identify the multiple uses of criterion- and norm-referenced test scores

16) communicate test results information effectively to such groups as parents, students, other school administrators, and board members.

These competencies are an ambitious list to be mastered in any one training program. But the list provides a starting point for planning training activities and support for better testing in Wyoming in the future. Wyoming is unique with respect to its population size and density, its rurality, its ethnically homogeneous population, and the independence of its local school districts. Wyoming public school students score far above national norms on nationally standardized tests. This unique set of characteristics must be taken into consideration in seeking to support districts wishing to include criterion-referenced testing in their testing programs.

List of References

Popham, W. J. & Hambleton, R. K. (1990). Can you pass the test on testing? Principal, 69(3), 38-39.

Weber, Ed (1989). Instructional Management System. Campbell County School District, Gillette, Wyoming.

Sheinker, Alan (August, 1990). Sweetwater School District Number One's Comprehensive Assessment Program. Paper presented at the University of Wyoming Trustee's Symposium, August 2-3, 1990, Laramie, Wyoming.

Jaeger, Richard M. (1987). Two decades of revolution in educational measurement? Educational Measurement: Issues and Practice, 6(4), 6-14.

Table 1.

1. Who is responsible for the test coordination in your district?

TITLE	FREQUENCY	PERCENT
Superintendent	9	17.6
Counselor	8	15.7
Curriculum Director	8	15.7
Assistant Superintendent	7	13.7
Principal(s)	4	7.8
Special Services Director(s)	4	7.8
Director of Instruction	2	3.9
Testing Coordinator	2	3.9
Administrative Assistant	1	2.0
Coordinator of Instr. Serv.	1	2.0
Director of Pupil Services	1	2.0
Educational Diagnostician	1	2.0
Instructional Support Leader	1	2.0
Monitor of Basic Skills	1	2.0
Chapter I Reading Teacher	1	2.0
	51	

Table 2.

What percent of the test coordinator's position is directly related to overseeing the district's testing program?

PERCENT OF POSITION	FREQUENCY	PERCENT
0-3	11	23.9
5	14	30.4
10	9	19.6
14-50	10	21.7
90-100	2	4.3
100	1	

Table 3.

If you are currently developing or using criterion-referenced testing, in which of the following stages are you?

	planning the tests		**developing** objectives		**writing** items		**using** tests		**modifying** tests	
	f	%	f	%	f	%	f	%	f	%
math	11	21.6	12	23.5	8	15.7	13	25.5	10	19.6
reading	11	21.6	10	19.6	3	5.9	11	21.6	8	15.7
other lang. arts	8	15.7	10	19.6	4	7.8	10	19.6	7	13.7
science	8	16.7	13	22.2	5	9.8	9	17.6	5	9.8
social studies	12	23.5	11	21.6	5	9.8	7	13.7	4	7.8

Table 4.

How great a need is there in your district for professional development in the area of the training in the development, validation, and utilization of criterion-referenced tests?

NEED	FREQUENCY	PERCENT
None	4	7.8
Slight	4	7.8
Moderate	14	27.4
Great	19	37.2
Extreme	8	15.7

Table 5.

If there is a need for professional development, what kind would be most desirable and feasible in your district? (check all that apply)

KIND OF PROFESSIONAL DEVELOPMENT	FREQUENCY	PERCENT
Individual on-site consulting with the district	38	74.5
Professional development seminars	30	58.8
Extension course work through UW	26	51.0
A summer institute	23	45.1
Computer support	15	29.4
Individual telephone consulting with the district	10	19.6
Course work at UW	6	11.8
Other	5	9.8

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 29, 1991